DOCUMENT RESUME

ED 067 670 CS 200 146

AUTHOR	Tucker, Shirley A.; And Others
TITLE	The Effects of Word Segmentation Schemes on Spelling Acquisition and Retention.
PUB DATE	7 Apr 72
NOTE	23p.; Paper presented at the Annual Meeting of the American Educational Research Assn. (Chicago, Ill., April, 1972)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Analysis of Variance; Computer Assisted Instruction; *Educational Research; Grade 3; Grade 4; Grade 5; Paired Associate Learning; *Retention; *Spelling
IDENTIFIERS	Bell Telephone Communication System; PDP 7 PDP 9 Computer System; Westinghouse CROW

ABSTRACT

This experiment was conducted to examine word presentation routines to determine their effectiveness in spelling drills. The design included segmentation or focal unit (letter, chunk, whole word), audio (audio, no audio), and word type (List 1--pronounceable chunks, List 2--pronounced in running speech). Subjects were 48 students in Grades 3, 4, and 5 spelling classes. Computer-assisted spelling drills were utilized for the experiment, in which misspelled words were given one of the six treatments. Retention tests were given at two weeks and six weeks after acquisition. Analysis of variance on learning rates and retention revealed that List 2 words were acquired faster than List 1 words and no variables reached significance on retention tests. T-tests, computed for acquisition of words in single sessions versus acquisition in more than one session, showed that 40 percent were massed and 60 percent acquired in more than one session. Overall t's were significant on the first retention test, indicating that distributed practice was superior to massed; however, on the second test, words were not differentially retained. Analyses of misspellings showed there were no differences in the percentages of misspellings involving errors within the four categories (additions, substitutions, omissions, and reversals). Substitutions were the most frequent error. Latency analyses indicate that correct responses have shorter mean word latencies than incorrect responses. (JF)

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED 067670

To be presented at AERA Session C-44
April 7, 1972
By Shirley A. Tucker

FILMED FROM BEST AVAILABLE COPY

THE EFFECTS OF WORD SEGMENTATION SCHEMES ON SPELLING ACQUISITION AND RETENTION.

Shirley A. Tucker and Karen K. Block
Learning Research and Development Center, University of Pittsburgh

Patricia A. Butler[1]
Southwest Regional Laboratory

The learning of a list of spelling words has been said to resemble the learning of a list of paired-associate (PA) items (Fishman, Keller & Atkinson, 1968; Knutson, 1967). The comparison provides a heuristic model for research into the design of spelling drills, in that variables known to influence PA learning might usefully be investigated in spelling drills. In first proposing this approach, Fishman et al. (1968) investigated the effects of massed versus distributed practice (i.e., practice on a word occurred within one day or on three alternating days) and demonstrated that massed practice resulted in a higher proportion correct during acquisition, but inferior retention. Knutson (1965) investigated repetition conditions whereby an incorrectly spelled word was requested again immediately after the misspelling, or requested after four intervening items, or not requested again during a session. Results generally showed within-session repetitions to be superior to no repetition, but

[1]The authors wish to express their appreciation to Richard Fellers of the Office of Measurement and Evaluation, University of Pittsburgh, for assistance in preparing the data for the NYBMUL computer program (Jeremy D. Finn, Version 4, June 18, 1969).

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
Shirley A. Tucker
Karen K. Block
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER"

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
Patricia A. Butler
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER"

CS 200 146

there were some interactions with spelling ability. In both of these computer-controlled studies, the presentation routine for each spelling word was the same: an audio system presented the words; if a spelling was correct, "C" was printed; if it was incorrect, "X" was printed, followed by the correct spelling. The purpose of the present experiment is to examine several word presentation routines to determine their relative effectiveness in spelling drills. The variable was examined in the context of a computer-assisted spelling program whose design incorporated Knutson's (1967) findings. (See Figure 1.) The program was tailored to fit an existing spelling curriculum managed under a system of individualized instruction.

In reviewing various ways to display a spelling word after it has been misspelled, it seemed that segmenting a word into smaller parts might be useful, in that acquisition could proceed "by parts." The part learning conditions could be compared to a whole word, or nonsegmented presentation as employed by Fishman et al. and Knutson to determine the relative effectiveness of segmentation.

The segmentation schemes used were selected to vary in "part" size, i.e., in the number of letters belonging to a group, to be applicable to a large corpus of polysyllabic words, and to differ in the way the segments related to the sounds of a word. While many segmentation schemes qualify under these criteria, a fourth criterion was imposed: the schemes should be analogous to schemes used in the teaching of spelling or reading. With this, two segmentation schemes were defined to be compared to a

whole word treatment: one whose segments defined pronounceable parts of a word (e.g., hol-i-day) and a second which segmented a word into single letters (h-o-l-i-d-a-y). The segmentations were accomplished when a word was spelled incorrectly; the word was re-presented on the terminal screen below the misspelling and the word was segmented into parts: three spaces occurred between the chunks, or one space occurred between the letters, or the whole word was displayed with no spaces. The chunk and whole word methods of displaying spelling words have been used in classroom experiments in spelling (Wolfe & Breed, 1922; Greene, 1923, Horn, 1969) with contradictory results. The single letter has been a focal unit in reading in the traditional phonic approach.

The segmentation scheme which produced pronounceable chunks used syllabication rules defined and modified by Burmeister (1968, Pp. 86, 87) and demonstrated to have high utility with words above the primary level in difficulty. A subset of the rules used in the present study included determining the number of syllables through a count of vowels and vowel combinations, requiring structural syllabication first (i.e., separating roots from each other and from affixes), and then performing phonic syllabication.

It was hypothesized that to the extent that each of the segmentation schemes was differentially compatible with the spelling learning process, then they should have a differential effect on acquisition, and possibly retention.

If either scheme enhances the learning or retention process, then this should be evident in its superiority relative to the whole word treatment. The segmentation schemes might also result in different response organizations in the manner defined by Johnson's (1970) analysis of Chunking in P-A learning. These organizations should be evident in the kinds of structural errors found in misspellings emitted in spelling recall. Using Johnson's analysis as an analogy, if the first letter of a chunk is recalled, the probability that the entire chunk is spelled correctly should be close to 1.00. In addition, isolating the middle chunk of a spelling word via the spatial display may effect a redistribution of attention during acquisition (cf. using color cues, underlining, etc.). If so, then the serial position curve for spelling errors for words learned via a chunking scheme should be flatter relative to the curve usually found for whole words (Jass and Gillooly, 1972). To the extent that the letter segmentation promotes unorganized learning of the response letters in a spelling, then the serial position curve should approach a monatonic decreasing function, to approximate curves found in P-A learning for responses consisting of letter clusters with low interletter association.

A second variable included in the study was audio enhancement. When a segmentation was displayed it either was or was not accompanied by audio pronunciation of the segments, or of the whole word. In the Chunk group, each chunk was pronounced as it was displayed, and the pronunciation of the chunk was as consistent as possible with its pronunciation in the whole word. In the Letter group, the letter name was said as each letter was displayed.

The audio enhancement was included at both levels with each segmentation scheme, to determine if it enhanced the effects of any segmentation scheme.

In addition to the segmentation and audio variables, each student learned two kinds of words: those in which every pronounceable chunk is heard in the pronunciation of the word in running speech (List 2 words: e.g., hol-i-day) in contrast to those in which a medial chunk is missing, (List 1 words: e.g., business whose chunks for spelling are bus-i-ness). The list variable was added to determine whether the special treatment given to the unsounded chunk in the chunking segmentation which "sounds out" and/or visually isolates the unsounded chunk, would aid in spelling recall, relative to the other two treatments which did not treat it differentially. Fourteen words were matched on number of letters, number of chunks, Thorndike-Lorge frequency, fourth grade difficulty rating on the Iowa Scale, and grade placement in extant spelling curricula.

Method

Design

The design contained two between-subject factors: segmentation or focal unit (letter, chunk, whole word) and audio (audio, no audio) and one within factor: word type (List 1 and List 2).

Subjects

Forty-eight students from the third, fourth and fifth grade spelling classes of a suburban elementary school were assigned randomly to the

six experimental conditions. Classroom and achievement level were counterbalanced for each cell of eight subjects.

Apparatus

An inter-linked PDP 7/PDP 9 computer system located at the Learning Research and Development Center utilized the Bell Telephone communication system to control student terminals located in a separate room in the elementary school. Each student station was equipped with a Datapoint (consisting of a cathode ray tube (CRT) and a response keyboard) and a set of earphones. A rapid random-access audio system, the Westinghouse CROW, was used. From one to three students were run simultaneously.

Procedure

The experiment was conducted in the context of a drill and practice spelling program. Students heard a word, read its sentence context on the CRT screen, and then spelled the word by typing it on a keyboard. If the word was spelled incorrectly, it received one of the six experimental treatments. The student studied the word as appropriately segmented and signaled when ready for an immediate retest of the same word. A second incorrect spelling was followed by the treatment and a fixed study time of four seconds. (See Figure 1.)

The two types of words were selected randomly from matched lists of fourteen words each until the student has missed five words from each list or a total of ten words, or until the lists were exhausted. Each word was repeated after an average lag of about six other words. The criterion for acquisition was three correct spellings in a row for each

word presented. However, only words misspelled at some point during acquisition were used in assessing the effects of the experimental variables. The final pool receiving experimental treatment ranged from a list split of 6 to 3 to the optimal 5 and 5. On the average, acquisition required three 20-minute sessions, each of which was separated by one or two days.

Retention

Retention testing of words receiving experimental treatment occurred two times: two weeks and six weeks after acquisition. The word was pronounced, a sentence displayed, and the student typed the spelling. After testing was completed, students received feedback and copied words missed onto a study sheet.

Results

Analyses of Variance on Learning Rates and Retention

Analyses of the inverse of the average number of trials to criterion (the transformation made the variance more homogeneous) revealed a strong main effect of word type. List 2 words (in which each chunk is pronounced in running speech) were acquired significantly faster than List 1 words. ($F = 12.8282$, $p < .0009$). The focal by list interaction approached significance ($F = 2.8619$, $p < .0684$; see Figure 2) indicating that the relative difficulty of the two lists had the greatest effect for the Letter group and the least effect for the Word group. No other main effect or interaction reached significance. A second ANOVA on acquisition data based on the inverse of average number of errors also revealed a list effect ($F = 8.5926$, $p < .0055$). However, no other effects approached significance.

Analyses of the proportion of words correct on the retention tests revealed that none of the experimental variables reached significance on either retention test. Only one effect approached significance: the main effect for focal unit on Retention Test One ($F = 2.7775$, $p < .0739$; see Figure 3). The two word types were not differentially retained. Also, the proportions of correct spellings recalled on the two retention tests were not substantially different (.63 and .67) yielding an average proportion correct of .65. Thus, the overall level of retention was stable over a long period of time.

Massed versus Distributed Comparisons.

Data were analyzed comparing massed versus distributed practice effects to the Fishman et al. program. Their program involved a fixed number of presentations for each word, whereas we continued trials to a criterion of three consecutively correct responses. Since our program was error-contingent, in that words were repeatedly practiced until they reached criterion, the number of sessions required for the acquisition of a word varied from one to three or more sessions. T-tests were computed for words acquired in one session versus those acquired in more than one session collapsing over the experimental variables. (See Table 1.) About 40 percent of the words were massed and 60 percent were acquired in more than one session.

For the first retention test, overall t's were significant indicating distributed practice was superior to massed acquisition. For the second retention test, the massed versus distributed words were not differentially retained.

The overall proportions correct for both massed and distributed words are slightly but consistently higher on both retention tests than the proportions observed by Fishman et al. Moreover, our retention intervals were longer (i.e., two and six weeks cf. ten and twenty days) and our data are based on "harder" words, i.e., words which were misspelled during acquisition. In contrast, their retention proportions contain 31 and 25 percent correct words on initial presentations. Note also that the massed versus distributed variables failed to influence retention assessed six weeks after acquisition. This is an interesting result and should be investigated in a planned experiment.

Insert Table 1 about here.

Analyses of Misspellings

The effects of the three focal units (or segmentation schemes) were investigated for four types of errors: additions, substitutions, omissions and reversals of adjacent letters (See Table 2). Generally, there were no differences in the percentage of misspellings involving the errors within each category using a χ^2 test assuming equal probability. Substitutions constituted the most frequent type of error for all three groups occurring in 51-55 percent of the misspellings.

Insert Table 2 about here.

To investigate the occurrence of missing chunks from List 1 words (e.g., bus-ness for bus-i-ness, choc-late for choc-o-late), misspellings were read aloud by three independent readers and scored for medial chunk omission in the pronunciation of the word. Segmentation treatments failed to influence the proportion of omitted medial chunks.

Serial position effect curves were plotted for words of equivalent length. Result support the established finding that errors are more probable in the middle of the word with initial and terminal letters correct. Figure 4 depicts the curves for the three focal groups for seven letter words and is representative of the data for other word lengths and other focal group comparisons. The three focal groups produce equivalent curves.

Misspellings on both retention tests were analyzed by adopting Johnson's (1970) analysis of the chunking of consonant strings in P-A learning. Obviously, Johnson's unpronounceable letter strings which carefully maintain low inter-letter associations are unlike the letter strings involved in the spelling task. However, it is interesting to compare the differences between the two tasks using Johnson's analyses. Johnson's theory is most appropriate for predicting response termination at chunk boundaries resulting from imperfect recall (stop-TEP's). We observed less than a dozen of these in over three hundred misspellings. Only three were attributed to the Chunk group; these responses did terminate at chunk boundaries. All three stop-TEP's found in the Word group, and one of the four from the Letter group occurred within a chunk as defined in our segmentation treatment for the Chunk group.

While Johnson finds the probability of letters correctly recalled given that the preceding letter is correctly recalled to be a monotonic decreasing function of the position of the letter in the string, we found serial position effects. Figure 5a is a graph of this curve for all word lengths (7-11 letters) and three focal groups combined. Figures 5b, 5c and 5d break out the curve for 7, 8 and 9 letter words found in the Letter group and are representative of the data from the other groups. Note that the "dip" occurs one position later as the number of letters in the word increases. Possibly the chunks of letters acquired together in learning spellings are more contingent upon the number of letters in the word than upon chunks organized by sounds. Data to support this conjecture require that dips are the same across words with different numbers of chunks. Unfortunately, our sample of words was not adequate to provide a test of this conjecture.

The adaptation for scoring spelling errors as shown in Figures 5a-5d involves scanning from left to right in the misspelling and identifying the first occurrence of a correct letter and proceeding forward. If a letter is not present, the scan returns to the occurrence of the last identification and proceeds forward. When the same letter occurs in two non-adjacent positions in the correct spelling and occurs only once in the misspelling, the letter is scored as an instance of the position that occurs within the longer sequence of correct letters. (For example, the "i" in SURSPRING is marked as the second "i" of "surprising"). Only those letters marked as present in the misspelling are used in determining whether the next letter (or the prior letter) in the misspelling is correct or not. Thus, on Figures 5a-5d the number of correct occurrences of Letter N is given below the position of Letter N+1.

Johnson's model predicts that letters chunked together tend to be recalled (or not recalled) together and the probability the total chunk is recalled is a function of the number of letters in the target chunk and the number of chunks in the sequence and is independent of the number of letters in subsequent chunks. We found that given that the first letter of the chunk is correct, the probability of the entire chunk being correct increased as chunk size increased. (See Figure 6.) Moreover, for chunk size two, the first letter was incorrect as often as it was correct. Therefore, 50 percent of the misspellings could not be scored. However, the serial position effect confounds these results since one and two letter chunks occur most frequently in the middle of the word. For words with only three and four letter chunks, (i.e., "controlling" chunk size) a serial position effect also occurred.

Figure 7 plots the proportion of total letters correctly recalled in order as a function of focal group and also by chunk size within the Chunk group. The experimental focal treatment produced no differential effects in such letter recall. Chunks of three and four letters yielded superior recall to chunks of one or two letters. The second result might be expected if we abandon Johnson's framework and again consider the errors as spelling errors. Vowels are notedly more difficult to spell than consonants. This holds true for all plots on the graph (one letter chunks contain only vowels) and most noticeably for two letter chunks.

Latency Analyses

The latency of correct and incorrect responses on the retention test were also analyzed. Using mean word latency over all groups combined

and on both retention tests, our results confirm Knutson's (1967) finding that correct responses have shorter mean word latencies than incorrect responses. Knutson also reported significant differences based on latencies of the first letter of the response during five acquisition runs. Using the same measures, we support his results for our first retention test (i.e., two weeks after acquisition). However, first letter latencies are not significantly different for correct and incorrect responses on our second retention test given six weeks after acquisition, although the mean word latencies do remain significantly different.

Discussion

With the exception of the effect of word type during acquisition, none of the experimental variables demonstrated clear-cut differences on any of the analyses described. Segmentation schemes with or without audio enhancement had no significantly differential effects on either acquisition or retention. Words that were harder to acquire were not differentially retained. Thus, these schemes are not differentially compatible with the processes of spelling learning for children at the intermediate level in the spelling curriculum. The response organization processes involved in the spelling task detected here through analyses of types of errors, serial position effects, and an adaptation of Johnson's chunking analyses appear to be relatively independent of segmentation schemes as realized in this experiment. Just as the ANOVA's revealed a lack of differentiation by focal group of the number of correct words retained, a second analysis revealed similarities in the number of correct letters recalled in order within the misspelled words.

In comparing results across the two retention tests, it is important to note that for all words receiving experimental treatment, the proportions of correct responses were stable from two to six weeks after acquisition. However, six weeks after acquisition, the differential effects of massed versus distributed practice during acquisition and the first letter latency of correct versus incorrect responses on Retention Test Two were no longer significant.

In conclusion, attempts to facilitate the spelling learning process using part-learning segmentation of the varities described here in spelling drills generally are not differential from nonsegmented or whole word treatments for our population of Ss. Thus, these segmentation schemes are independent of how a spelling is learned in such drills. These schemes provide no differential information to aid the learning process since they do not diagnose and, thus, are not adaptive to individual student's errors and they appear not to influence response organization and letter recall differentially. Future experimentation involving computer-assisted spelling drills might make word displays following errors a differential function of those errors to provide the student with the information that needs to be incorporated into a mneumonic to help him overcome his "hard spots." For example, a common error was to spell the vowel found in first position of the ending syllable of preference with an ⟨a⟩. Isolating this error spot in a display might help the development of a mneumonic such as "preference has all e's." Isolating the student's "hard spots" may help his spelling recall more than information given by segmentations which is the same for all errors.

REFERENCES

Burmeister, L. E. Selected word analysis generalizations for a group approach to corrective reading in the secondary school. Reading Research Quarterly, 1968, 4, 71-95.

Fishman, E. J., Keller, L., & Atkinson, R. C. Massed versus distributed practice in computerized spelling drills. Journal of Educational Psychology, 1968, 59, 290-296.

Green, H. A. Syllabication as a factor in learning to spell. Journal of Educational Research, 1923, 8, 208-219.

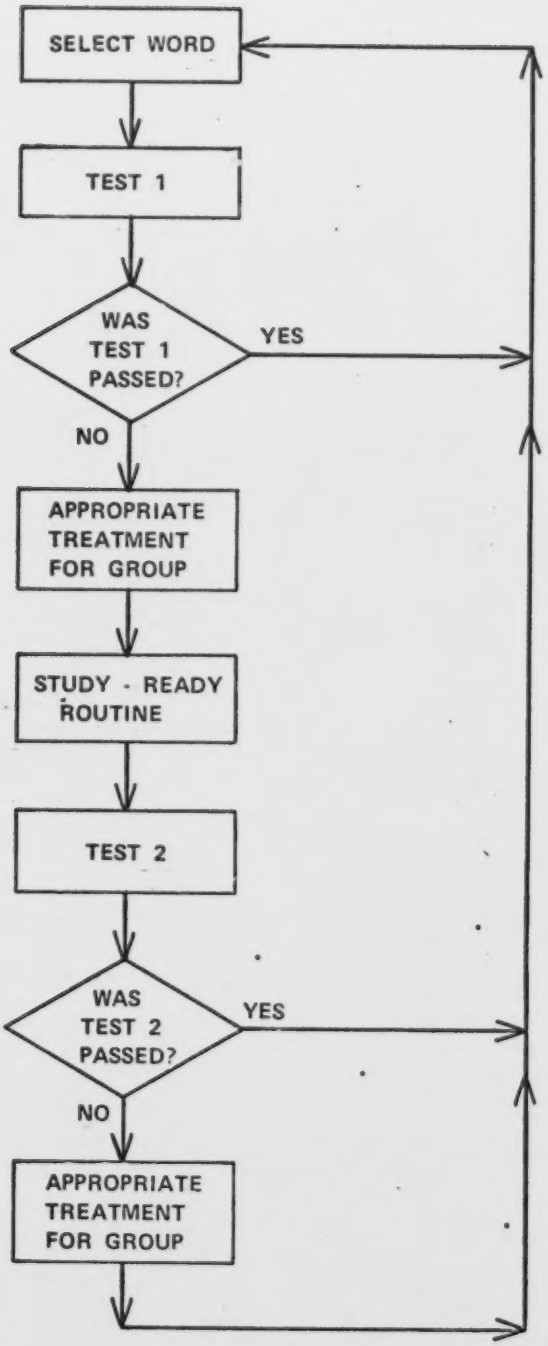
Horn, T. D. Spelling in Elsel (Ed.), Encyclopedia of Educational Research, Fourth Edition, 1969. Pp. 1282-1299.

Jass, A. F., & Gillooly, W. B. The Von Restorff effect on the number and distribution of spelling errors. Paper to be presented at the meeting of the American Educational Research Association, Chicago, April, 1972.

Johnson, N. F. The role of chunking and organization in the process of recall. In G. H. Bower (Ed.), The psychology of learning and motivation, Vol. IV. New York: Academic Press, 1970. Pp. 171-247

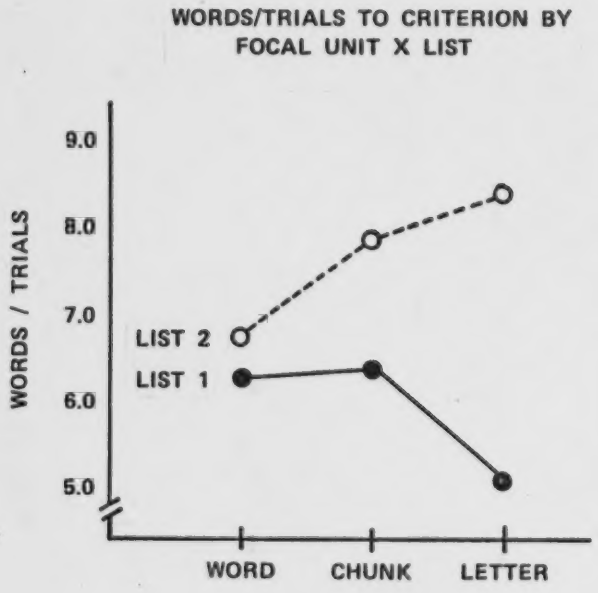
Knutson, J. M. Spelling drills using a computer-assisted instructional system. Technical Report No. 112, Institute for Mathematical Studies in the Social Sciences. Stanford: Stanford University, 1967.

Wolfe, H. A., & Breed, F. S. An experimental study of syllabication in spelling. School and Society, 1922, 15, 616-622.



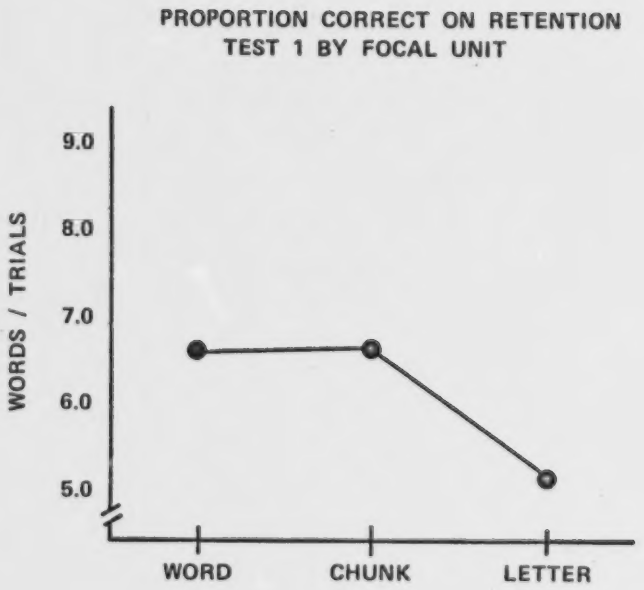
WORD SEGMENTATION EXPERIMENT: MAINLINE FLOW CHART

Figure 1



(Interaction Approaches Significance F= 2.8619 p<.0684)

Figure 2

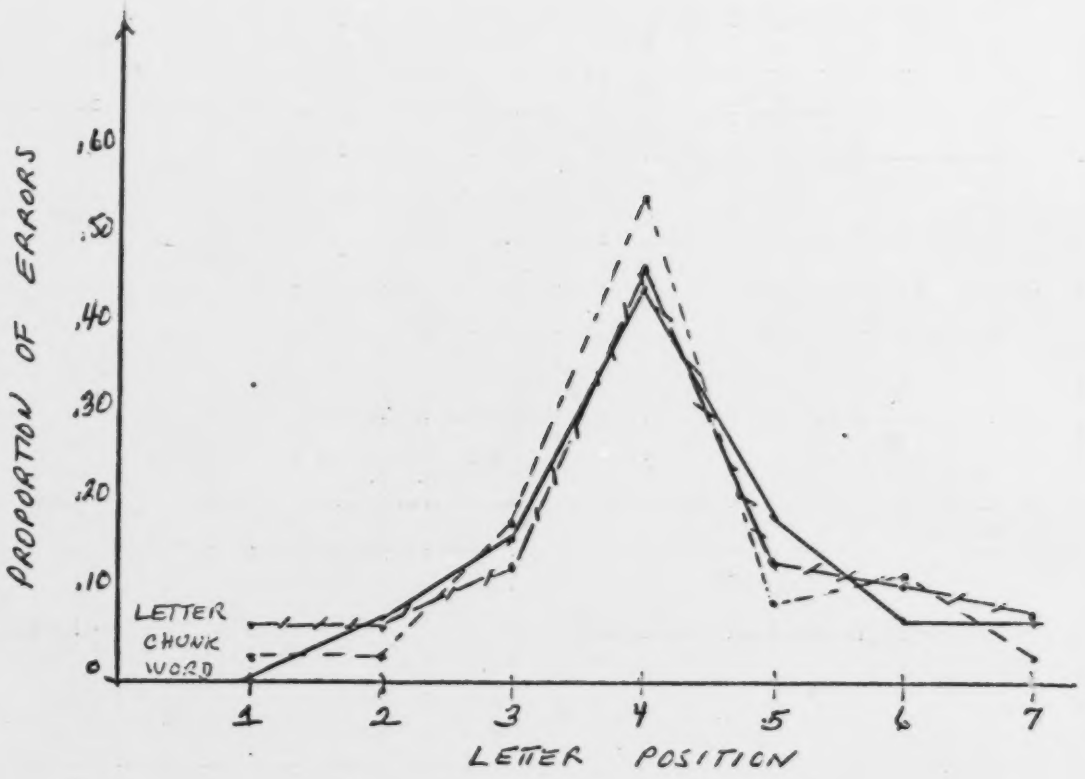


(Main Effect Approaches Significance F= 2.7775 p<.0737)

Figure 3

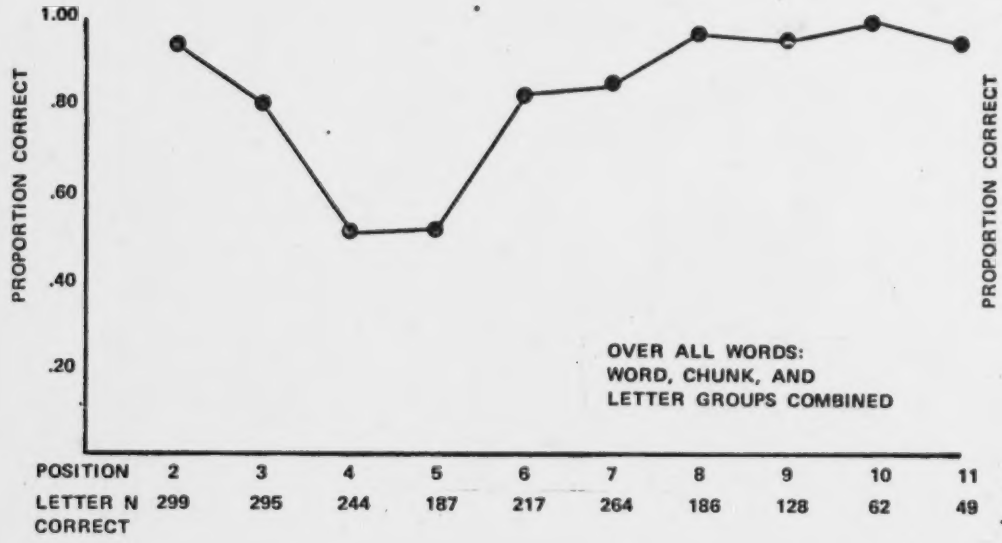
FIGURE 4

PROPORTION OF ERRORS BY LETTER POSITION FOR 7-LETTER WORDS



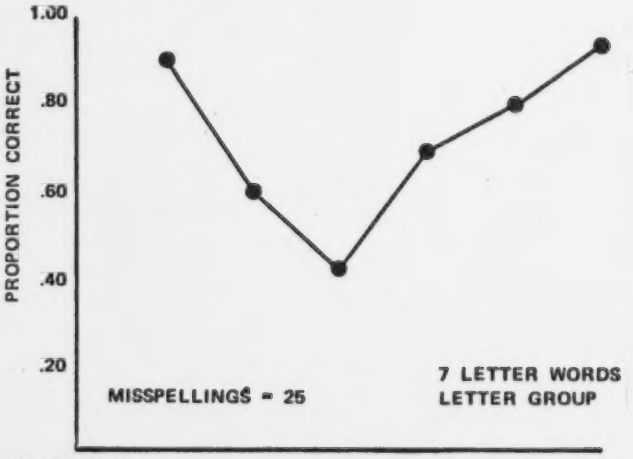
NUMBER OF MISSPELLINGS	
WORD	26
CHUNK	29
LETTER	25
TOTAL	80

GIVEN LETTER N IS CORRECT, LETTER N + 1 PROPORTION CORRECT BY POSITION



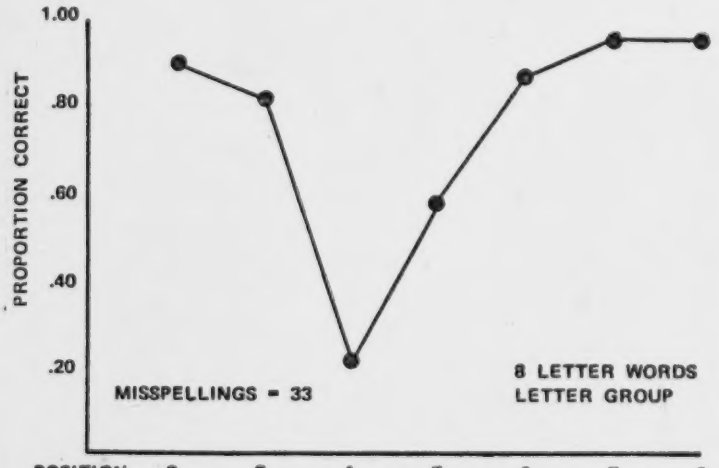
POSITION	2	3	4	5	6	7	8	9	10	11
LETTER N CORRECT	299	295	244	187	217	264	186	128	62	49

Figure 5a



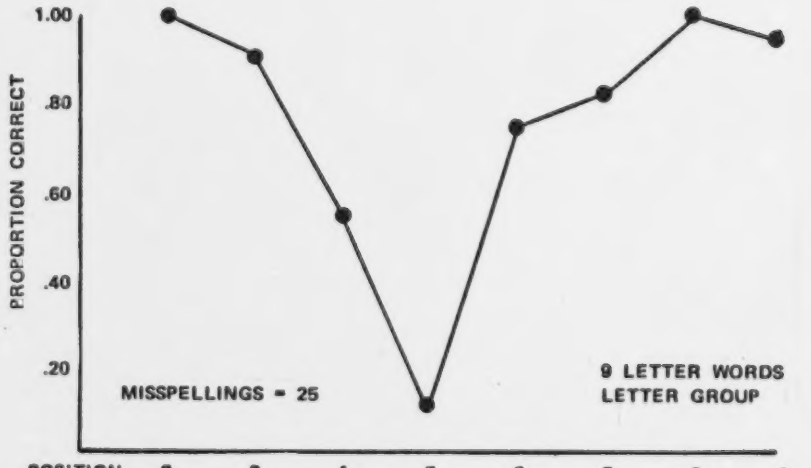
POSITION	2	3	4	5	6	7
LETTER N CORRECT	23	25	16	13	20	22

Figure 5b



POSITION	2	3	4	5	6	7	8
LETTER N CORRECT	33	31	32	12	30	31	31

Figure 5c.

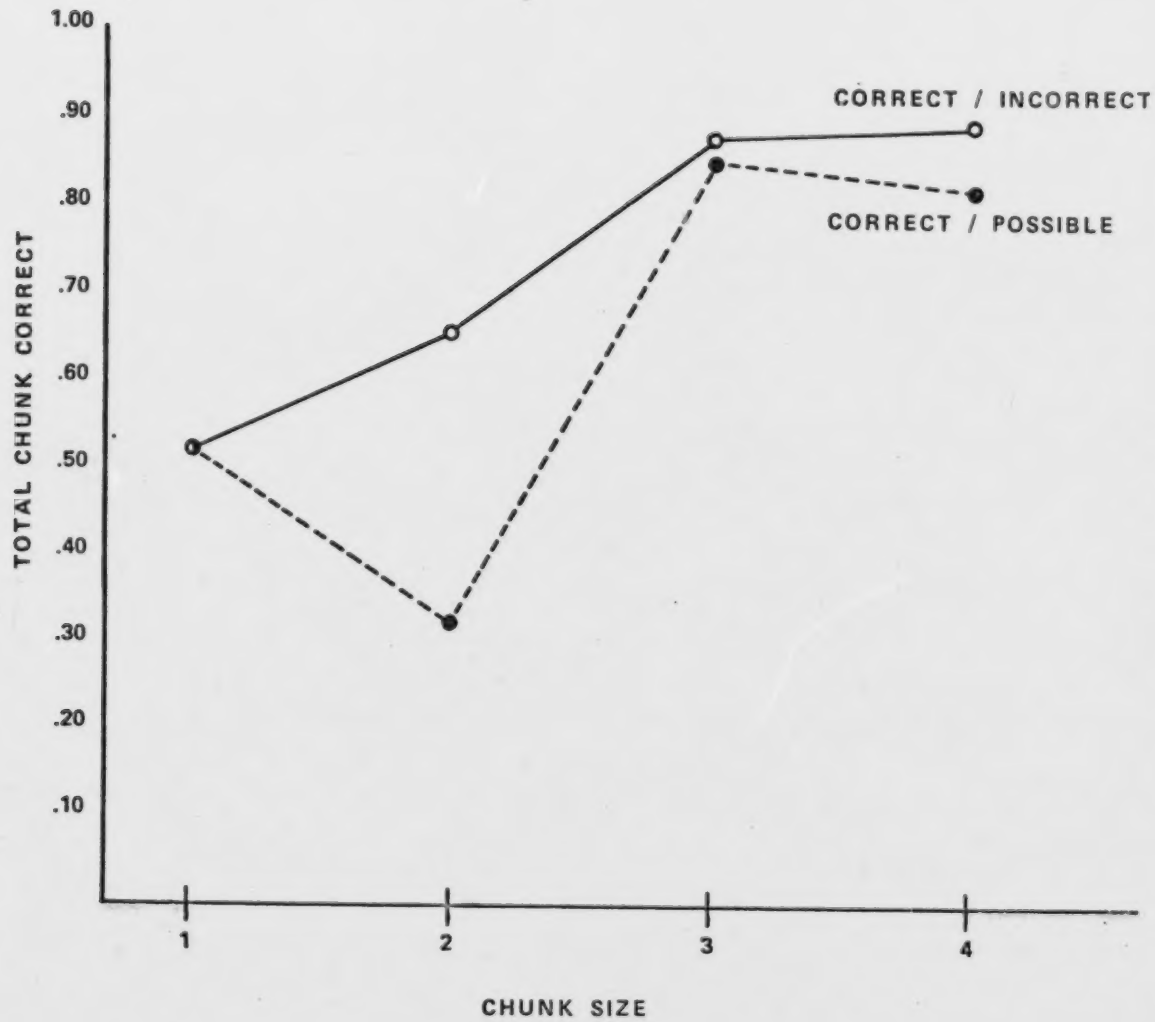


POSITION	2	3	4	5	6	7	8	9
LETTER N CORRECT	25	25	23	18	9	24	24	24

Figure 5d

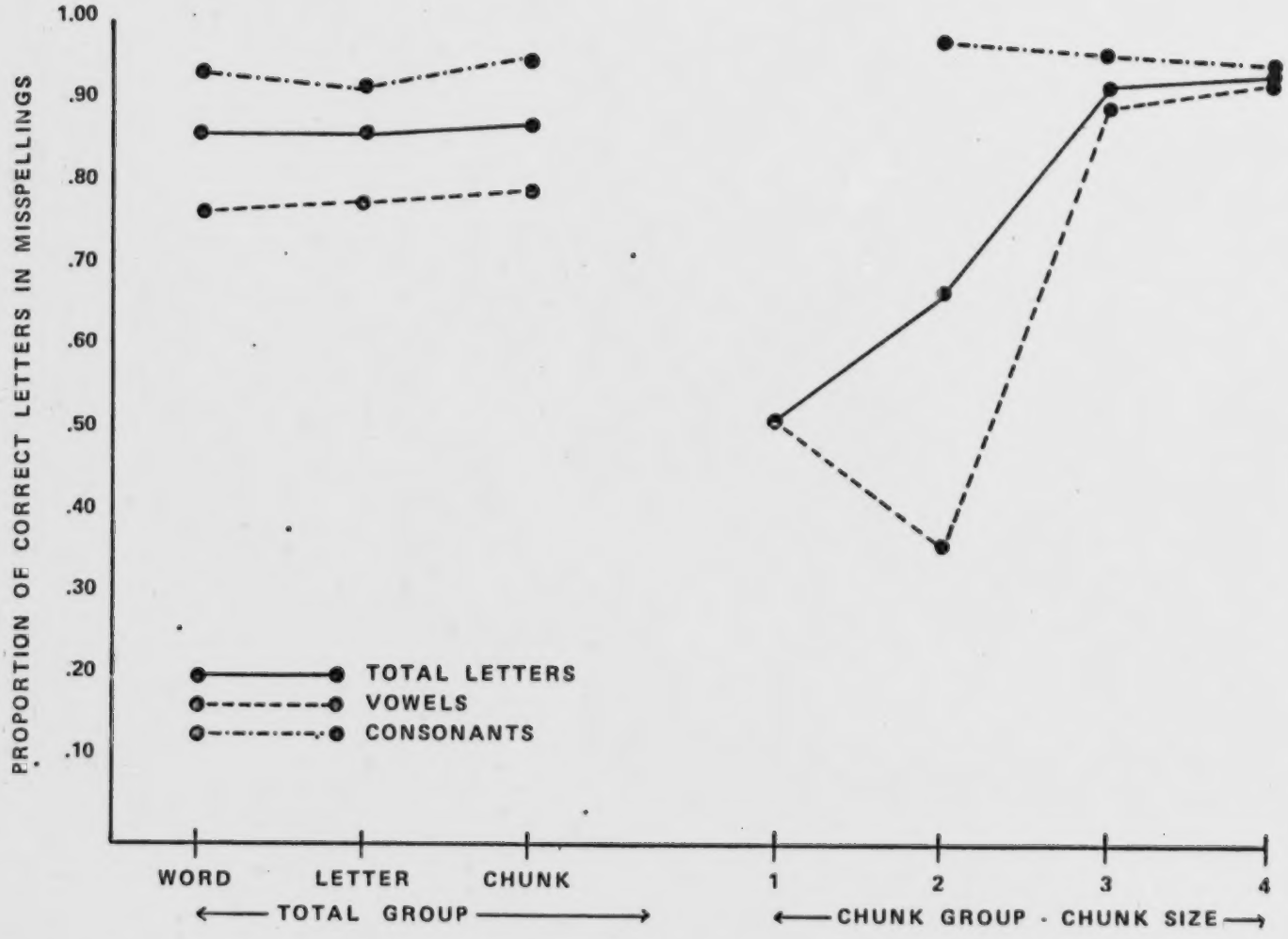
TOTAL CHUNK CORRECT AS A FUNCTION OF CHUNK SIZE

(Given that the first letter is correct)



NOT SCORED	0	23/46	.5000	3/136	.0220	9/64	.1406

Figure 6



N MISSPELLINGS	97	112	285	39	46	136	64
N TOTAL LETTERS	805	961	791	39	88	408	256
N VOWELS	327	429	324	39	44	155	86
N CONSONANTS	478	532	467	0	44	253	170

Figure 7

TABLE 1

Overall T-tests for Massed vs. Distributed Acquisition and Proportion Correct on Retention Tests

	One Session		More than One Session		t value	p <
	N words	Proportion Correct	N words	Proportion Correct		
Retention Test 1 (2 weeks after acquisition)	139	.5411	201	.7233	-3.7807	.005
Retention Test 2 (6 weeks after acquisition)	131	.6122	209	.6820	-1.1480	. 15

Fishman, Keller, Atkinson Data Proportion Correct

	Massed	Distributed
Test 1 (10 days after acquisition)	.51	.58
Test 2 (20 days after acquisition)	.55	.61

TABLE 2

Percent of Error Category to Total Misspelled Words
For Each Focal Unit

Focal Unit	Number of Misspellings	% Additions	% Omissions	% Substitutions	% Reversals
Word	100	19.	47.	57.	10.
Chunk	102	29.41	39.21	52.94	0.98
Letter	113	17.69	57.52	51.32	5.30
Average Percent per Category		22.03	47.91	53.75	5.42